Journal of Hydrology (2006) 327, 564–577

available at www.sciencedirect.com

SCIENCE DIRECT®

journal homepage: www.elsevier.com/locate/jhydrol

An advanced regularization methodology for use in watershed model calibration

John Doherty [a,b], Brian E. Skahill [c,*]

[a] *Department of Civil Engineering, University of Queensland, St Lucia, Australia*
[b] *Watermark Numerical Computing, Brisbane, Australia*
[c] *Coastal and Hydraulics Laboratory, US Army Engineer Research and Development Center, Waterways Experiment Station, 3909 Halls Ferry Road, Vicksburg, MS 39180, USA*

Received 21 March 2005; received in revised form 27 November 2005; accepted 30 November 2005

KEYWORDS
Regularization;
Calibration;
Watershed modeling

Summary A calibration methodology based on an efficient and stable mathematical regularization scheme is described. This scheme is a variant of so-called ''Tikhonov regularization'' in which the parameter estimation process is formulated as a constrained minimization problem. Use of the methodology eliminates the need for a modeler to formulate a parsimonious inverse problem in which a handful of parameters are designated for estimation prior to initiating the calibration process. Instead, the level of parameter parsimony required to achieve a stable solution to the inverse problem is determined by the inversion algorithm itself. Where parameters, or combinations of parameters, cannot be uniquely estimated, they are provided with values, or assigned relationships with other parameters, that are decreed to be realistic by the modeler. Conversely, where the information content of a calibration dataset is sufficient to allow estimates to be made of the values of many parameters, the making of such estimates is not precluded by ''preemptive parsimonizing'' ahead of the calibration process.

While Tikhonov schemes are very attractive and hence widely used, problems with numerical stability can sometimes arise because the strength with which regularization constraints are applied throughout the regularized inversion process cannot be guaranteed to exactly complement inadequacies in the information content of a given calibration dataset. A new technique overcomes this problem by allowing relative regularization weights to be estimated as parameters through the calibration process itself. The technique is applied to the simultaneous calibration of five subwatershed models, and it is demonstrated that the new scheme results in a more efficient inversion, and better enforcement of regularization constraints than traditional Tikhonov regularization methodologies. Moreover, it is argued that a joint calibration exercise of this type results in a more meaningful set of parameters than can be achieved by individual subwatershed model calibration.

* Corresponding author. Tel.: +1 601 634 3441; fax: +1 601 634 4208.
E-mail address: brian.e.skahill@erdc.usace.army.mil (B.E. Skahill).

0022-1694/$ - see front matter
doi:10.1016/j.jhydrol.2005.11.058

Introduction

''Regularization'' is a mathematical term that, in its broadest sense, refers to any measure that is taken to ensure that a stable solution is obtained to an otherwise ill-posed inverse problem. In traditional calibration practice, this is achieved through adherence to the so-called ''principle of parsimony'' in which parameters requiring adjustment through the calibration process are reduced to a number for which a unique estimate can be obtained for each. If calibration is computer-assisted, then prior to initiating execution of the parameter estimation software package, the modeler selects those parameters that he/she wishes to estimate (normally on the basis of anticipated higher sensitivity of model outputs to these parameters) and holds other parameters fixed at ''sensible values''.

Often this method works well. However, problems associated with this approach include the following:

1. It is not always possible to know ahead of the parameter estimation process how many parameters can be estimated. If too few are selected for estimation it may not be possible to obtain a good fit between model outputs and field measurements. If too many are selected, the parameter estimation process may suffer from numerical instability (which may seriously hamper its performance in maximizing model-to-measurement fit) and/or result in the estimation of a set of parameter values that lack credibility.
2. If, notwithstanding steps taken to reduce the number of parameters requiring estimation, large variability of one or a number of parameters leads to only minor changes in the goodness of fit between model outputs and field measurements (as a result of parameter insensitivity, a high level of measurement noise, or both), unique estimation of those parameters becomes problematical. This often results in meaningless values being assigned to them through the parameter estimation process.
3. Individual parameter sensitivities are not the sole determinator of what is estimable and what is not. Situations are often encountered where model outputs have a low sensitivity to certain parameters collectively, but can be very sensitive to the same parameters individually. This is the phenomenon of ''parameter correlation''.
4. Traditional approaches to calibration are not well suited to the solution of complex inverse problems, such as those involving simultaneous calibration of multiple models, where the estimation of useful values for otherwise nonuniquely estimable parameters may be assisted through the provision of trans-model parameter relationships, from which a departure will only be tolerated if supported by the calibration dataset. Nor can they be readily adapted to make best use of ''outside information'' on parameter values forthcoming from such disparate sources as soils mapping, satellite imaging, or simply the modeler's expert judgment.

The last point is particularly important, for exclusion from the calibration process of important information on watershed hydraulic properties reduces the ability of that process to estimate parameters that are ''robust'' in the sense of allowing the model to make useable predictions of watershed behavior under possibly different climatic and/or land management conditions than those that prevailed during the calibration time period, and of transporting knowledge gained from the calibration process to neighboring unguaged watersheds. Hence considerable attention has been devoted to incorporating knowledge of watershed properties (including a stochastic description of their spatial variability) into the calibration process (see, for example, Koren and Kuchment, 1971; Koren, 1993; Schaake et al., 1996), and of developing regional relationships between model parameters and measurable watershed characteristics (see, for example, Magette et al., 1976; Campbell and Bates, 2001; Yokoo et al., 2001).

All of the problems outlined above can be overcome through the use of parameter estimation algorithms that allow mathematical regularization to be implemented as part of the parameter estimation process itself. The result is a stable solution to the inverse problem (regardless of how ill-posed it is), and avoidance of the deleterious effects of numerical instability on both the parameter estimation process itself, and on the outcomes of that process, namely the set of estimated parameter values. A well-designed regularization algorithm, like its manual counterpart, achieves numerical stability by re-formulating the inverse problem in a way that recognizes the level of parsimony that is necessary to attain a stable solution to that problem. However, this ''parsimonizing'' is undertaken in the context of a specific calibration dataset, allowing numerical stability to be achieved without compromising model-to-measurement fit any more than is deemed necessary by the modeler.

(In using the term ''numerical stability'', we refer to the fact that the performance of some optimization methods, particularly the Gauss–Marquardt–Levenberg method described below, can deteriorate badly in the face of ill-posedness of the inverse problem. As will be discussed below, these methods are very efficient; however because their efficiency relies on matrix inversion as an integral part of the determination of a suitable parameter upgrade path, their ability to improve parameter estimates is seriously degraded if that matrix is singular as a result of parameter nonuniqueness. This expresses itself in an inability to improve model-to-measurement fit beyond a certain level that is easily recognized as far from optimal, accompanied often by oscillating estimates of parameter values through the iterative process through which these are calculated for nonlinear models.)

Doherty and Johnston (2003) demonstrated the use of regularized inversion in watershed model calibration. They employed a ''Tikhonov'' or ''constrained minimization'' scheme in which values assigned to estimated parameters were permitted to deviate from those defining a user-supplied ''preferred system parameterization'' only to the extent necessary for a desired level of model-to-measurement fit to be achieved. If the ''preferred system parameterization'' is defined wisely, this approach guarantees that reasonable parameter values are obtained, no matter how many parameters are estimated through the regularized inversion process (Tikhonov and Arsenin, 1977; Engl et al., 1996). This is because, with the calibration dataset thus supplemented by information pertaining directly to the

model parameters themselves, its information content is thereby sufficient to allow estimation of all such parameters.

This paper introduces an improvement to this scheme and demonstrates its use in calibrating a Hydrological Simulation Program Fortran (HSPF) (Bicknell et al., 2001) hydrologic model. The methodology is available through the PEST (Doherty, 2005) package which, together with its Surface Water Utility Suite (Doherty, 2003b), is freely available to the public; see below. (''PEST'' is an acronym for ''Parameter ESTimation''.)

Before describing this method it should be pointed out that, whether regularization is undertaken prior to commencing the calibration process through manual parsimonizing, or as an integral part of this process through mathematical means, the resulting parameter set cannot represent the hydraulic property detail of the real world, but instead represents average hydraulic properties over all or part of the simulated system. Furthermore, the amount of averaging required to achieve numerical stability rises as the information content of the calibration dataset decreases (see, for example, Backus and Gilbert, 1967; Menke, 1984). Where model predictions are then made using these averaged parameters, these predictions have the potential to be in error, especially where they depend on system property details that are not represented in the model. Moore and Doherty (2006) show that the magnitude of this error can be considerable. Hence post-calibration analysis of model predictive error variance should become a routine adjunct to model calibration and deployment; see Moore and Doherty (2005) for details of a methodology through which such an analysis can be undertaken.

Theory

Gauss–Marquardt–Levenberg parameter estimation

Let the action of a model under calibration conditions be described by the model operator $\mathscr{M}$ that maps m-dimensional parameter space to the space of the n observations that are available for use in the calibration process. Let the m-dimensional vector $\mathbf{p}$ represent model parameters and the n-dimensional vector $\mathbf{h}$ represent observations. In many instances of watershed hydrologic model calibration these observations will represent stream discharges which have been ''processed'' in some way in order to achieve homoscedascity, and statistical independence of measurement ''noise''. The former is often achieved through a Box–Cox transformation (Box and Cox, 1964), while the latter is often attempted through fitting residuals to an ARMA model, often as part of the parameter estimation process itself (Box and Jenkins, 1976; Kuczera, 1983). The observations $\mathbf{h}$ can be comprised of a single observation type, multiple observation types, and/or a single observation type processed in different ways in order to ensure that the information content associated with different aspects of the calibration dataset exercise sufficient influence in the estimation of a final set of model parameters (Madsen, 2000; Boyle et al., 2000; Doherty and Johnston, 2003).

Model calibration seeks to minimize some measure of model-to-measurement misfit encapsulated in a ''measurement objective function'', herein designated as Φ_m. In the present instance this is defined as

$$\Phi_m = [\mathscr{M}(\mathbf{p}) - \mathbf{h}]^t \mathbf{Q} [\mathscr{M}(\mathbf{p}) - \mathbf{h}] \quad (1)$$

where $\mathbf{Q}$ is a ''weight matrix'' which, in the context of watershed model calibration where n is large, is mostly comprised of diagonal elements only. Ideally, each diagonal element of $\mathbf{Q}$ is proportional to the inverse of the squared potential error associated with the corresponding processed measurement.

Where $\mathbf{p}$ is estimable (i.e. where minimization of Φ_m results in a unique parameter set), it is calculated as

$$\mathbf{p} - \mathbf{p}_0 = (\mathbf{X}^t \mathbf{Q} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{Q} (\mathbf{h} - \mathbf{h}_0) \quad (2)$$

where $\mathbf{X}$ is the model Jacobian matrix, each row of which is comprised of the derivatives (i.e. sensitivities) of a particular model output (for which there is a corresponding field measurement) with respect to all elements of $\mathbf{p}$. These sensitivities are calculated at current parameter values, represented by $\mathbf{p}_0$, for which corresponding model outputs are $\mathbf{h}_0$. Where the model is nonlinear, $\mathbf{p}$ calculated through Eq. (2) is not optimal (i.e. it does not minimize Φ_m) unless $\mathbf{p}_0$ is close to optimal. Hence, after Eq. (2) is used to calculate an improved parameter set, a new set of sensitivities (i.e. $\mathbf{X}$) is calculated on the basis of the new parameter set, and the process is repeated until convergence to the objective function minimum is achieved.

In practice, the $\mathbf{X}^t\mathbf{Q}\mathbf{X}$ matrix of Eq. (2) is supplemented by addition of a diagonal term – the so-called ''Marquardt lambda''. Thus, Eq. (2) becomes

$$\mathbf{p} - \mathbf{p}_0 = (\mathbf{X}^t \mathbf{Q} \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^t \mathbf{Q} (\mathbf{h} - \mathbf{h}_0) \quad (3)$$

Normally λ is adjusted during each iteration of the parameter estimation process such that its current value results in maximum parameter improvement during that iteration. When λ is high it is easily shown that the direction of parameter improvement is the negative of the gradient of Φ_m and under these conditions Eq. (3) becomes equivalent to the ''steepest descent'' method of parameter estimation. While this method can result in rapid parameter improvement when parameters are far from optimal, its performance is disappointing in the vicinity of the objective function minimum, especially where that minimum occupies a long valley in parameter space as a result of excessive parameter correlation or insensitivity. In these circumstances ''hemstitching'' is likely to occur, where successive parameter improvements result in oscillations across the objective function valley, which is never actually penetrated (Doherty, 2005). Hence, ideally λ should commence the parameter estimation process with a moderate value, and then be reduced as the process progresses. However, if $\mathbf{X}^t\mathbf{Q}\mathbf{X}$ is ill-conditioned, reducing the value of λ will incur numerical instability as $\mathbf{X}^t\mathbf{Q}\mathbf{X} + \lambda\mathbf{I}$ of Eq. (3) is inverted. Hence, the Marquardt lambda has a secondary role, this being that of a de facto regularization device, with its value often being raised in order to prevent instability in the calculation of the parameter upgrade vector $\mathbf{p} - \mathbf{p}_0$. However, while the use of a high Marquardt lambda can prevent a relatively ill-posed parameter estimation problem from foun-

dering, it achieves this at a cost in efficiency, for parameter upgrades become smaller at higher values of λ as an inspection of Eq. (3) suggests. Furthermore, as stated above, the ability of the calibration process to penetrate an elongate valley in parameter space may be severely compromised.

The predisposition of a matrix to stable inversion is often measured by its ''condition number''. High condition numbers result in amplification of numerical noise during the inversion process (Conte and de Boor, 1972) while low condition numbers indicate that inversion should be possible with little numerical difficulty. In general, condition numbers for $\mathbf{X}^t\mathbf{Q}\mathbf{X}$ greater than about 10^4 are to be avoided when using software such as PEST, for at this level the numerical noise incurred through finite difference-based derivatives calculation for filling of the $\mathbf{X}$ matrix is amplified to the extent that parameter upgrades may lack integrity. While a raised Marquardt lambda can often rescue such a damaged process from total failure as described above, efficiency of the parameter estimation process is likely to be seriously degraded.

Another problem that can be encountered when parameter estimation is accomplished by iterative calculation of $\mathbf{p} - \mathbf{p}_0$ using (3), is that this process can converge to a parameter set $\mathbf{p}$ that corresponds to a local, rather than the global, minimum of the objective function. ''Gradient methods'', such as the Gauss–Marquardt–Levenberg method described above, that rely on equations such as (3) have been criticized for this reason, and so-called ''global search'' methods such as SCE-UA (Duan et al., 1992) are often used instead. While a well-designed and robust global search method can indeed be guaranteed to minimize the objective function in spite of the existence of local minima, such robustness comes at a price, this being the high number of model runs that is normally required for completion of the parameter estimation process. To make matters worse, the number of model runs increases dramatically as the number of parameters requiring estimation increases. Use of Eq. (3), on the other hand, is very run-efficient. Fortunately, its propensity to find local minima can be mitigated through the use of schemes such as that encapsulated in the PD_MS2 software package described by Doherty (2003b) which combines the efficiency of gradient methods with the benefits of introducing a small degree of randomness to the parameter estimation process, together with an ability to ''learn from past mistakes''. In addition, Eq. (3) can be enhanced by the inclusion of a regularization term (much more powerful than the Marquardt lambda as will be described shortly) that greatly increases the propensity for robust and efficient behavior when the dimension m of $\mathbf{p}$ is large, and the shape of the objective function surface in parameter space becomes a valley (or series of valleys) rather than a bowl (or series of bowls).

It is the opinion of the authors that all inversion methods have strengths and weaknesses. It is not therefore our intention to recommend one over the other for universal application. It is our desire, however, to demonstrate one of the strengths of the Gauss–Marquardt–Levenberg approach, this being its ability to readily incorporate regularization into the inversion process, and to demonstrate a means whereby this can be achieved in a more numerically stable manner when applied to watershed model calibration than has been done in the past.

Regularized inversion

Conceptually, singularity or near-singularity of $\mathbf{X}^t\mathbf{Q}\mathbf{X}$ (as occurs when large numbers of parameters require estimation and/or when the information content of the calibration dataset with respect to estimated parameters is poor) can be remedied through the addition of extra ''observations'' to the parameter estimation process which pertain directly to the parameters requiring estimation. For example, it may be ''observed'' that each parameter is equal to a certain, user-supplied value; presumably this value will have been chosen to be realistic in terms of the system property which the parameter represents. Alternatively (or as well), it may be ''observed'' that certain pairs of parameters are equal, or have values which observe a certain ratio or difference.

Let these ''regularization relationships'' be represented by the operator $\mathscr{Z}$ acting on the parameter set $\mathbf{p}$, and let the ''observed'' values of these relationships be represented by $\mathbf{j}$. Then the regularization relationships (also referred to as ''regularization constraints'' herein) can be represented by the equation:

$$\mathscr{Z}(\mathbf{p}) = \mathbf{j} \tag{4a}$$

the linearized form of which is

$$\mathbf{Z}\mathbf{p} = \mathbf{j} \tag{4b}$$

where $\mathbf{Z}$ is the Jacobian of the $\mathscr{Z}$ operator. Note that, as is discussed below, it is not essential that (4a) and (4b) be exactly observed, only that they be observed to the maximum extent possible in calibrating the model.

If the regularization constraints are given sufficient weight in comparison with the observation weights encapsulated in $\mathbf{Q}$, a well-posed inverse problem will have been formulated. Mathematically, this problem is then iteratively solved for the parameters $\mathbf{p}$ using the equation:

$$\mathbf{p} - \mathbf{p}_0 = (\mathbf{X}^t\mathbf{Q}\mathbf{X} + \beta^2\mathbf{Z}^t\mathbf{S}\mathbf{Z} + \lambda\mathbf{I})^{-1}(\mathbf{X}^t\mathbf{Q}[\mathbf{h} - \mathbf{h}_0] + \beta^2\mathbf{Z}^t\mathbf{S}[\mathbf{j} - \mathbf{j}_0]) \tag{5}$$

In Eq. (5) $\mathbf{j}_0$ represents the right side of (4a) when current parameter values $\mathbf{p}_0$ are substituted for $\mathbf{p}$ in this equation. $\mathbf{S}$ is a ''relative weight matrix'' assigned to the regularization observations $\mathbf{j}$; it has the same role for regularization observations as $\mathbf{Q}$ does for field observations. All of the relative regularization weights encapsulated in $\mathbf{S}$ are multiplied by a ''regularization weight factor'' β^2 in Eq. (5) prior to calculation of $\mathbf{p} - \mathbf{p}_0$. $\mathbf{S}$ is supplied by the user. In many instances it will consist simply of a set of weights applied individually to the regularization observations $\mathbf{j}$, such that those with greater weight are more rigidly enforced; in this case S is a diagonal matrix.

Selection of an appropriate value for β^2 is critical. If its value is too high the parameter estimation process will ignore the measurement dataset $\mathbf{h}$ in favor of fitting the regularization observations $\mathbf{j}$. If it is too small, the regularization observations will not endow the parameter estimation process with the numerical stability which it needs in order to obtain estimates for the parameters $\mathbf{p}$. Alternatively, the assignment of a value to β^2 can be viewed as a mechanism for trading parameter reasonableness against goodness of fit. On the assumption that parameter value

reasonableness underpins the definition of regularization observations **j**, and that closer adherence to these regularization conditions therefore results in more reasonable values for model parameters, the use of a high value for β^2 should lead to a high degree of parameter reasonableness in the calibrated model. However, it also results in the assignment of values to model parameters in isolation from measurements of the state of the system whose physical properties they purport to represent. In contrast, if β^2 is set too low, insufficient recognition is given to the noise associated with measurements of system state; consequential ''overfitting'' can then introduce errors to parameter estimates (and to predictions that depend on them) as reasonableness is abandoned in ruthless pursuit of a good fit between model outputs and field measurements. Fortunately, as will be discussed shortly, a mechanism is available for selection of a value for β^2 which, in many calibration contexts, can be considered optimal for that context.

Eq. (5) can be shown to constitute a constrained minimization problem (deGroote-Hedlin and Constable, 1990; Doherty, 2003a) in which a ''regularization objective function'' Φ_r defined as

$$\Phi_r = [Z(\mathbf{p}) - \mathbf{j}]^t \mathbf{S} [Z(\mathbf{p}) - \mathbf{j}] \tag{6}$$

is minimized subject to the constraint that Φ_m of Eq. (1) rises no higher than a user-specified value, referred to herein as the ''target measurement objective function''. Thus the user informs the regularized inversion process of the level of model-to-measurement misfit required; this process then enforces the regularization constraints defined through Eq. (4a) to the maximum extent that it can by minimizing Φ_r subject to the constraint that Φ_m rises no higher than the target level. If the target measurement objective function cannot be achieved, the regularized inversion process simply minimizes Φ_m; however, where minimization of Φ_m would otherwise be an unstable process due to parameter nonuniqueness, stability of this process is maintained by seeking that set of parameters lying within the elongate Φ_m valley that also minimizes Φ_r. In either case, the regularization weight factor β^2 can be viewed as a Lagrange multiplier associated with the constrained minimization problem. In PEST it is re-calculated during every iteration of the regularized nonlinear parameter estimation process using a bisection algorithm based on local linearization of the constrained minimization problem about current parameter values. (For the linearized problem β^2 is calculated to ensure that Φ_m coincides exactly with its user-supplied target value.)

Note the continued inclusion of the Marquardt lambda in Eq. (5). In PEST its value is adjusted as needed from iteration to iteration as a practical measure to enhance optimization efficiency and to ensure stability of the parameter estimation process should $\mathbf{X}^t\mathbf{Q}\mathbf{X} + \beta^2\mathbf{Z}^t\mathbf{S}\mathbf{Z}$ become ill-conditioned through use of an inappropriately low value for β^2. This can occur where regularization constraints are poorly formulated, or where too a good a fit is sought between model outputs and field measurements, requiring that regularization constraints be abandoned in pursuit of this fit. Often it occurs for a combination of these reasons, where weights on some regularization constraints must be lowered for attainment of a good fit between model outputs and field measurements, but where the relaxation of regularization constraints then leads to unestimability of those model parameters whose estimation is not realized through attainment of this fit.

Formulation of the inverse problem as a constrained minimization problem through use of Eq. (5) allows many more parameters to be estimated than would otherwise be possible, thereby ensuring that maximum information is extracted from the calibration dataset. If the relationships of Eq. (4) are realistic, the fact that estimated parameters are such as to ensure minimal deviation from these relationships heightens the probability that estimated parameters will themselves be realistic. However, a practical problem that is often encountered when using the Tikhonov method is that the regularization weight matrix **S** must be supplied ahead of the regularized inversion process; furthermore, it is not adjusted through this process except for global multiplication by β^2. Ideally, individual regularization constraints described by the rows of Eq. (4) should be more strongly enforced where the information content of the calibration dataset is insufficient to require their contravention for the sake of obtaining an appropriate level of model-to-measurement fit. However because it is almost impossible to know ahead of the calibration process the extent to which this should occur for each of the different relationships encapsulated in $\mathscr{Z}$, it is often very difficult to supply an **S** matrix that is an appropriate complement to the current calibration dataset. This is especially the case where, as in the example presented below, the parameters upon which $\mathscr{Z}$ operates fall into a number of groups of very different type, and possess very different sensitivities to the observations comprising the calibration dataset. Thus a value for β^2 which may guarantee estimation of a sensible set of values for one parameter type (because it prevents overfitting on the one hand, or misfitting on the other hand through too strong an enforcement of regularization constraints) may be wholly inappropriate for the members of another parameter group, often resulting in unrealistic estimates for values of parameters comprising that group at best, and poor performance of the optimization package at worst because the condition number of the $(\mathbf{X}^t\mathbf{Q}\mathbf{X} + \beta^2\mathbf{Z}^t\mathbf{S}\mathbf{Z} + \lambda\mathbf{I})$ matrix of Eq. (5) then becomes too high for numerically stable inversion.

Nevertheless, especially where parameters are of the same or similar type (for example, if they characterize the spatial distribution of a particular hydraulic property over a model domain), Tikhonov regularization can be both effective and efficient as a calibration device. Hence it has been employed with great success over many years in many different disciplines, sometimes with ingenious design of the regularization operator $\mathscr{Z}$ for optimization of the ability of the inversion process to estimate parameter sets that are realistic in different modeling contexts. See, for example, Constable et al. (1987), Portniaguine and Zhdanov (1999) and Haber et al. (2000) in the geophysical context; Emsellem and de Marsily (1971), Skaggs and Kabala (1994), Bruckner et al. (1998), in the groundwater modeling context; and van Loon and Troch (2002) and Doherty and Johnston (2003) in the surface water modeling context.

Adaptive regularization

An ''adaptive regularization'' methodology is now presented which overcomes this problem in many modeling contexts. The set of regularization constraints described by Eq. (4) is subdivided into groups; if desired, each constraint can be assigned to its own group. The set of model parameters **p** is then supplemented by an additional parameter set $\mathbf{p}_r$, with one new parameter being defined for each new regularization group. Each such parameter is, in fact, the inverse of a group-specific regularization weight multiplier; this group-specific weight multiplier is applied in addition to the global weight multiplier β^2 depicted in Eq. (5), the latter being adjusted as part of the constrained minimization process as described above. Regularization constraints are then provided for the elements of $\mathbf{p}_r$ so that these too can be estimated as part of the regularized inversion process. Each such constraint comprises the ''observation'' that the respective element of $\mathbf{p}_r$ is zero.

The re-formulated regularized inversion problem remains a constrained minimization process, and thus still seeks to find a parameter set that either minimizes the measurement objective function Φ_m, or reduces it to a user-specified target level, while ensuring that the regularization objective function Φ_r is conditionally minimized. Because conditional minimization of the regularization objective function now requires maximization of weights assigned to individual or groups of regularization constraints, these weights are applied as strongly as possible, thereby maximizing the extent to which the corresponding regularization relationships encapsulated in Eq. (4) are adhered. However, with the calculation of the overall regularization weight factor β^2 by the constrained minimization process being such as to allow minimization of the target measurement objective function, or achievement of a user-specified target for this function, these regularization constraints are not so strongly enforced that model-to-measurement fit is compromised. Thus, the regularized inversion process itself ensures that the strength of enforcement of regularization constraints on parameter values or relationships complements the information content of the calibration dataset in relation to these parameters. As a result, regularization constraints are automatically applied more strongly where the attainment of a satisfactory level of model-to-measurement fit does not require otherwise, thus overcoming a disadvantage of the Tikhonov method. The outcome is a numerically stable regularized inversion process that achieves a desired level of model-to-measurement fit with impressive run economy, and that yields sensible values for model parameters.

Like all numerical strategies, this adaptive regularization methodology is more suitable for use in some contexts than in others. It is certainly not the only means by which numerical stability of a regularized inversion process can be achieved, for so-called ''subspace methods'' (Aster et al., 2005), and hybrid schemes such as ''SVD-Assist'' (Tonkin and Doherty, 2005) are very effective in this regard. However, use of the present methodology can be beneficial in those modeling contexts where the means by which numerical stability is achieved is just as important as the achievement of that stability itself. In general, where the necessity for parameters to observe key values or relationships to the maximum extent possible without compromising fit between model outputs and field measurements is a critical part of the calibration process, then the adaptive regularization methodology described herein will serve that calibration process well; such a case is demonstrated in the following section. However, the need to introduce extra parameters into the calibration process in order to guarantee enforcement of desired parameter relationships does place some restrictions on the method. Where such relationships fall into a relatively small number of distinct groups, and/or where the number of parameters requiring estimation is not such as to introduce vastly different levels of ''estimability'' between them (thus requiring the introduction of many new parameters in order to accommodate the differential strengths with which regularization constraints must be applied), the above method has proven very successful. However, where large numbers of parameters require estimation, and where differences in estimability between them are likely to cover a broad range, recourse to subspace methods becomes a necessity. Unfortunately, in this case, the guarantee of numerical stability that accompanies use of such methods is attained at the cost of loss of ability on the part of the modeler to insist on the observance of specified parameter relationships in attaining that stability.

An example

Model description

An HSPF hydrologic model that was developed as part of a total maximum daily load study (ENVVEST Regulatory Working Group, 2002) for the approximately 42 square kilometer Chico Creek watershed located in Kitsap County, Washington, USA was used for the purpose of demonstrating the benefits of adaptive regularization relative to traditional calibration methodologies. The HSPF model includes separate submodels for the drainage areas upstream of five streamflow gaging stations (Kitsap Creek, Wildcat Creek, Chico Creek Tributary at Taylor Road, Dickerson Creek, and Chico Creek mainstem) located within the watershed. The location of the Chico Creek Watershed in Kitsap County is depicted in Fig. 1.

The names and roles of model parameters selected for adjustment through the calibration process are provided in Table 1. Also listed are the bounds on these parameters imposed during the parameter estimation process, these being set in accordance with available guidance from, for example, USEPA (2000). Five instances of all but the first of the parameters listed in Table 1 required estimation, one instance for each subwatershed model. In contrast, the first adjustable model parameter type listed in Table 1, IMP, pertains to all five subwatersheds simultaneously. It possessed four instances however, one for each of four land use types occurring within the Chico Creek watershed; a preferred value was assigned to each instance as a regularization constraint (see Table 2b). Thus a total of 49 model parameters required estimation through the calibration process. In order to better accommodate scaling issues resulting from the use of different units for different parameters, and in an attempt to decrease the degree of nonlinearity of the

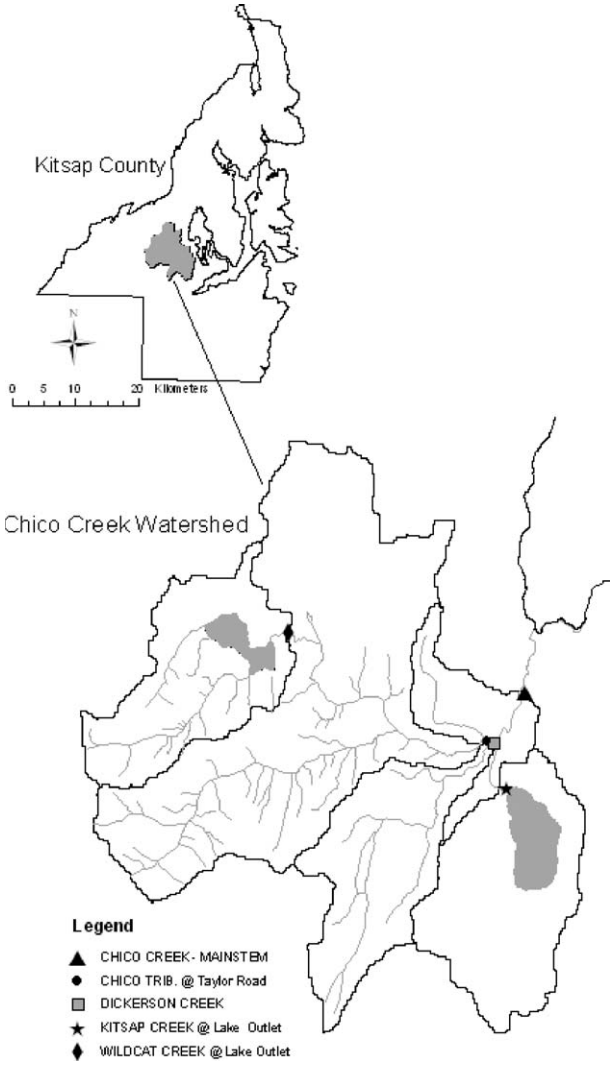

Figure 1 Location of the Chico Creek watershed in Kitsap County, Washington, USA.

parameter estimation problem, the logs of these parameters were estimated instead of their native values; past experience has demonstrated that greater efficiency and stability of the parameter estimation process can often be achieved through this means (Doherty and Johnston, 2003).

Simultaneous estimation of the parameters listed in Table 1 for the five different subwatersheds allows an important piece of information to be included in the parameter estimation process. Namely that, due to their geographical proximity and similarity of land use, soil type, and other geomorphic and anthropogenic conditions, parameter values employed in the different subwatershed models are not expected to be significantly different. To accommodate this condition, a series of regularization constraints effecting an assumed similarity condition across the subwatersheds was included in the regularized parameter estimation process. That is, respective log differences of identical parameter types between subwatersheds were ascribed an ''observed value'' of zero. The advantage of supplying such information through regularization constraints rather than through ''hardwired'' parameter equality is that the regularized inversion process has the option of introducing parameter differences if this is a requirement for obtaining a good fit between modeled and observed flows at each of the streamflow gaging stations. However, the constrained optimization algorithm which underpins the regularized inversion process guarantees that only the minimum amount of inter-parameter variability required to achieve this level of fit is introduced. Thus, subwatershed individuality is recognized at the same time as subwatershed similarity.

Mean daily discharge data associated with the five streamflow gaging stations located within the Chico Creek Watershed was available for the period 1st January 2001 to 31st December 2002, with some data gaps for each system. (The inadequacies of a limited-duration dataset as a basis for model calibration are freely acknowledged; use of the present dataset is justified by the fact that no other data were available. It should be noted, however, that these inadequacies do not detract from the role of the present paper in demonstrating a methodology for extracting as much information as possible from a dataset such as this — or from any other dataset.) Values for the 49 adjustable model parameters were estimated through simultaneous calibration against the mean daily discharge data at all five streamflow gauging stations, with mean daily flows transformed according to the equation (Box and Cox, 1964):

$$h_i = \ln(q_i + 0.01) \tag{7}$$

where h_i is the ''observation'' employed in the actual parameter estimation process (this being the ith element of $\mathbf{h}$ of Eq. (1)), and q_i is the corresponding measured mean daily flow. (As stated above, this type of transformation is one of a continuum of flow transformations often employed in the calibration of watershed models to promote homoscedascity of measurement noise; see, for example, Bates and Campbell, 2001.)

Results

Table 2a lists parameter values for each subwatershed model, estimated using the adaptive regularization scheme described above. Estimated impervious area percentages are listed in Table 2b, together with the preferred values for these parameters employed in the regularization constraints applied to them. In implementing the regularized inversion process, a very low target measurement objective function was set; hence PEST was asked to lower Φ_m of Eq. (1) as far as possible, thus reducing misfit between measured and observed flows to a minimum. It is apparent from Table 2a that optimal fitting of model outputs to daily flows could only be achieved through the assignment of different values to parameters of the same type in different subwatersheds. However, the adaptive regularization scheme employed in their estimation attempted to ensure that these differences were kept to a minimum. Fig. 2 shows the fit between the logs of modeled and observed flows at the Wildcat Creek streamflow gaging station; similar fits were obtained at the other streamflow gaging stations. The total measurement objective function (pertaining to all streamflow gages) achieved through this calibration exercise was 135.1.

Table 1 Parameters estimated in calibration of Chico Creek subwatershed models

Parameter name	Parameter function	Bounds imposed during calibration process
IMP	Percent effective impervious area	11—19% for med. dens. residential 19—32% for high dens. residential 51—98% for comm. and industrial 7—10% for acreage and rural residential (Alley and Veenhuis, 1983)
AGWETP	Fraction of ET taken from groundwater (after accounting for that taken from other sources)	0.0—0.2
AGWRC	Groundwater recession parameter	0.833—0.999 day^{-1}
DEEPFR	Fraction of groundwater inflow that goes to inactive groundwater	0.0—0.2
INFILT	Related to infiltration capacity of the soil	0.003—2.5 cm/h
INTFW	Interflow inflow parameter	1.0—10.0
IRC	Interflow recession parameter	0.30—0.85 day^{-1}
LZETP	Lower zone ET parameter — an index of the density of deep-rooted vegetation	0.1—0.9
LZSN	Lower zone nominal storage	5—38 cm
UZSN	Upper zone nominal storage	0.12—5 cm

PEST required a total of 1117 model runs to estimate values for the 49 parameters involved in the adaptive regularization inversion process. Each model run required about 20 seconds for completion on a 3 GHz Pentium 4 machine; thus the time required for completion of the entire adaptive regularization process was about 6 hours. (This could have been reduced dramatically through the use of Parallel PEST in conjunction with one or a number of other networked computers.)

The combined subwatershed parameter estimation process was repeated with equality-based regularization constraints replaced by hardwired parameter equality for all but the IMP parameters; thus 45 adjustable parameters were replaced by 9. However, the four impervious area parameters were estimated with the same regularization constraints as those described above, bringing the number of estimated parameters to a total of 13. An objective function of 349.5 was achieved through this process. Estimated parameter values are listed in Table 3, while measured and modeled flows for Wildcat Creek are compared in Fig. 3. It is apparent from the attained objective function, and from a comparison between Figs. 2 and 3, that flows in Chico Creek subwatersheds are better simulated when inter-subwatershed parameter variation is allowed. This is verified by an examination of Nash—Sutcliffe coefficients computed for the logs of modeled and observed flows at individual streamflow gaging stations, and collectively at all gaging stations, based on calibration fits obtained

Table 2a Estimated values for subwatershed model parameters for attainment of best fit at all Chico Creek subwatershed streamflow gaging stations, this corresponding to a measurement objective function of 135.1

Parameter	Kitsap Creek	Wildcat Creek	Chico Creek (Taylor Road)	Dickerson Creek	Chico Creek (mainstream)
AGWETP	2.08E—03	1.75E—03	1.55E—03	1.83E—03	1.92E—03
AGWRC	0.985	0.982	0.964	0.984	0.975
DEEPFR	9.00E—03	7.37E—03	1.26E—02	7.53E—03	1.18E—02
INFILT	0.36	0.11	0.091	0.12	0.19
INTFW	1.42	2.53	1.64	2.95	1.56
IRC	0.81	0.63	0.71	0.72	0.73
LZETP	0.28	0.41	0.57	0.12	0.59
LZSN	17.8	19.7	33.1	20.5	18.2
UZSN	3.94	3.45	5.08	4.75	2.82

Adaptive regularization was employed in the parameter estimation process.

Table 2b Estimated values for fractional impervious area parameters for attainment of best fit at all Chico Creek subwatershed streamflow gaging stations

Parameter	Regularization constraint	Estimated value
IMP1 (med. dens. residential)	0.15	0.16
IMP2 (high dens. residential)	0.23	0.20
IMP3 (comm. and industrial)	0.83	0.63
IMP4 (acreage and rural residential)	0.084	0.096

through adaptive regularization on the one hand, and hard-wired parameter equality on the other hand; see Table 4. The superiority of the fit obtained where inter-subwatershed parameter variability is allowed is readily apparent from this table.

Comparative performance with other regularization methods

Further PEST runs were undertaken in order to compare the performance of the adaptive regularization scheme discussed herein with that of other regularization methodologies.

In the first of these runs, simultaneous calibration of the five Chico Creek subwatersheds was repeated using an

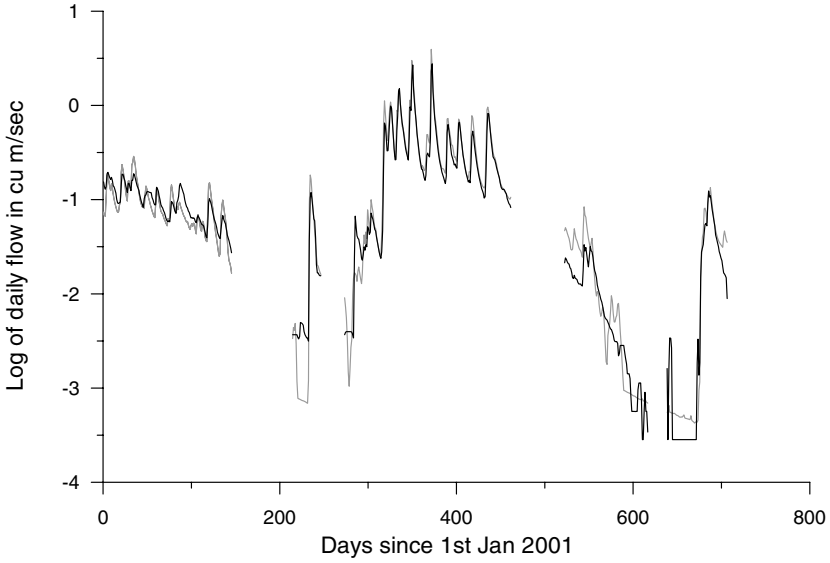

Figure 2a Observed (bold) and modeled (light) Wildcat Creek daily flows over the calibration period. Calibration achieved through simultaneous adaptive regularization.

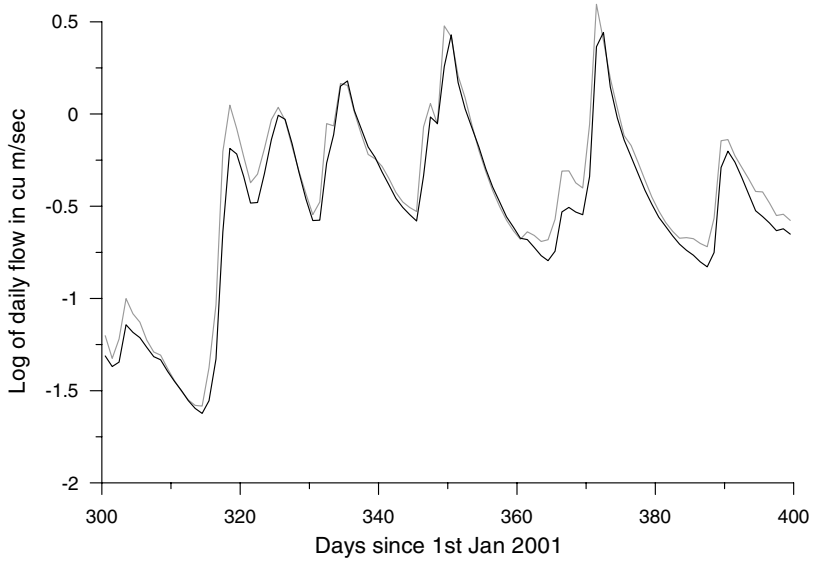

Figure 2b A magnified portion of Fig. 2.

Table 3 Estimated parameter values for all Chico Creek subwatersheds where parameter equality constraints are rigidly enforced

Parameter	Estimated value
IMP1	0.18
IMP2	0.23
IMP3	0.90
IMP4	0.10
AGWETP	1.15E−03
AGWRC	0.976
DEEPFR	1.03E−02
INFILT	0.25
INTFW	1.88
IRC	0.71
LZETP	0.38
LZSN	27.9
UZSN	2.52

identical Tikhonov scheme to that described above, but without the use of adaptive regularization to ensure maximal enforcement of regularization constraints. A similar measurement objective function and Nash—Sutcliffe fitting coefficients were obtained using this calibration methodology to those obtained using the adaptive regularization approach; model calculated flows were virtually identical to those depicted in Fig. 2. However, the condition number of the $\mathbf{X}^t\mathbf{Q}\mathbf{X} + \beta^2\mathbf{Z}^t\mathbf{S}\mathbf{Z}$ matrix of Eq. (5) for most iterations of the parameter estimation process was about 10^5, compared to about 10^3 when using adaptive regularization. Because of this, PEST found it necessary to employ much higher (by a factor of between 100 and 1000) values for the Marquardt lambda than was required for adaptive regularization. The result was slower convergence of the parameter estimation process, with over 1500 model runs being required for its completion.

Table 5 lists parameter values estimated though this process. It is readily apparent that inter-subwatershed variability between parameters of the same type is much greater for Table 5a than it is for Table 2a. Furthermore, a number of parameters listed in Table 5a are at their upper or lower bounds. This is also true of the IMP parameters depicted in Table 5b, three out of four of which are at their limits in defiance of the regularization constraints imposed on these parameters through which the preferred values listed in Table 2 were assigned to them.

The reason for the poorer numerical performance of the unenhanced Tikhonov scheme becomes obvious upon inspection of the weights ($\beta^2\mathbf{S}$ of Eq. (5)) applied to the regularization relationships comprising the $\mathbf{Z}$ matrix of Eq. (4) in formulation of the regularization objective function. For simultaneous calibration of the five Chico Creek subwatersheds most of the rows of $\mathbf{Z}$ consist of equality relationships between parameters of the same type in different subwatersheds; five such relationships (one for each subwatershed pair) exist for all but the IMP parameters. An extra four rows of $\mathbf{Z}$ are then employed for the assignment of IMP preferred values. Because none of the regularization relationships comprising the rows of $\mathbf{Z}$ are assumed to possess any statistical correlation, $\mathbf{S}$ was supplied as a diagonal matrix.

In implementation of an unenhanced Tikhonov scheme the weights comprising the elements of $\mathbf{S}$ are not adjusted relative to each other; rather they are modified solely through uniform multiplication by the regularization weight factor β^2, this being calculated to maximize inter-subwatershed parameter uniformity subject to the constraint that the measurement objective function rises no higher than a certain, user-supplied, level; re-calculation of β^2, based on a local linearization assumption, is undertaken during every iteration of the PEST nonlinear parameter estimation process. At the end of the unenhanced Tikhonov regularized inversion process through which parameters were simultaneously estimated for all Chico Creek subwatershed models,

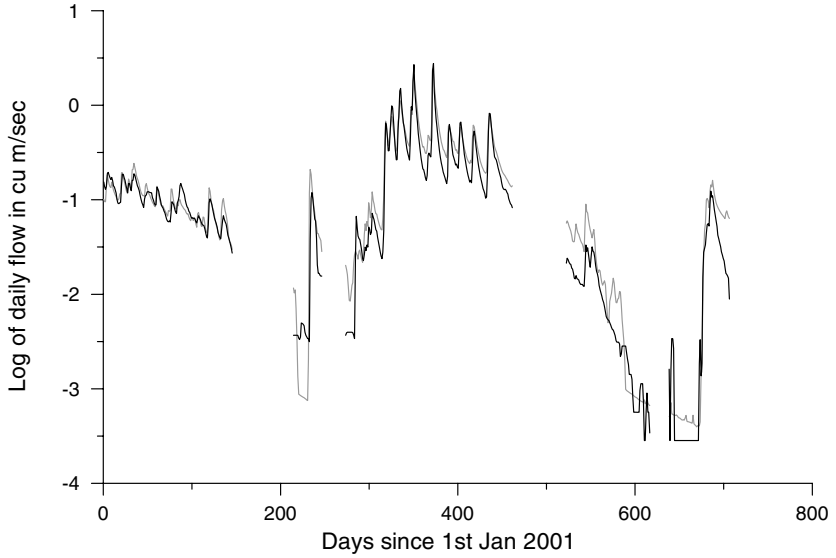

Figure 3a Observed (bold) and modeled (light) Wildcat Creek daily flows over the calibration period. Like parameters from different subwatersheds constrained by hardwired equality.

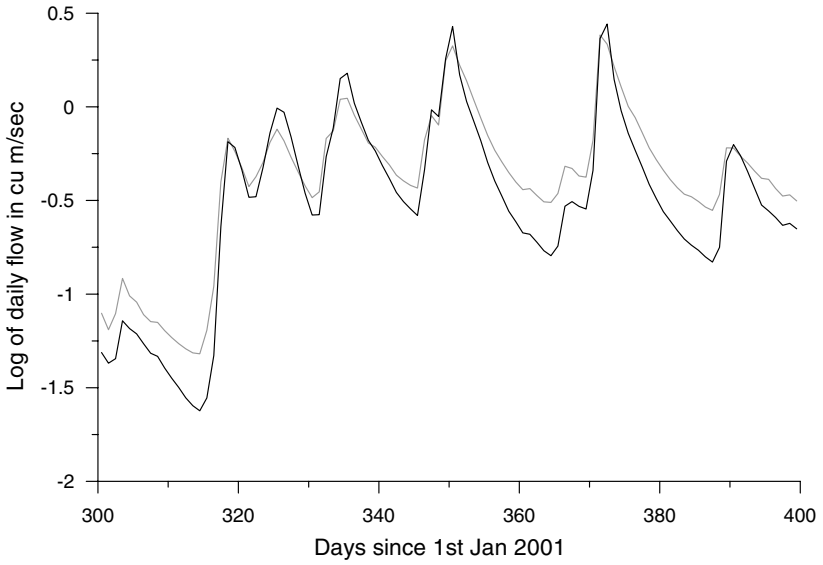

Figure 3b A magnified portion of Fig. 3a.

Table 4 Nash–Sutcliffe coefficients for log of daily flows based on simultaneous calibration through regularized inversion (column 2) and simultaneous calibration with hardwired parameter equality (column 3)

Streamflow gaging station	Adaptive regularization	Hardwired parameter equality
Kitsap Creek	0.768	0.336
Wildcat Creek	0.918	0.879
Chico Creek (Taylor Road)	0.888	0.675
Dickerson Creek	0.936	0.879
Chico Creek (mainstream)	0.952	0.916
All gaging stations	0.917	0.846

PEST-calculated values for all of the diagonal elements of **S** were 2.37E−2. The fact that they were all equal is a direct consequence of the fact that the originally supplied **S** matrix contained equal diagonal elements.

Regularization weights calculated by PEST through the adaptive regularization process leading to the parameter estimates provided in Table 2 were starkly different. In this case, multipliers for subgroups of the diagonal elements of **S** are actually estimated through the parameter estimation process under the constraint that they be as large as possible without compromising model-to-measurement misfit. Computation of weights in this manner not only allows enforcement of equality constraints to vary between parameter groups; it also reduces the potential for underestimation of weight multipliers, for less reliance is placed on

Table 5b Estimated values for fractional impervious area parameters for attainment of best fit at all Chico Creek subwatershed streamflow gaging stations using a traditional Tikhonov scheme

Parameter	Regularization constraint	Estimated value
IMP1	0.15	0.19
IMP2	0.23	0.32
IMP3	0.83	0.51
IMP4	0.084	0.07

Table 5a Estimated values for subwatershed model parameters for attainment of best fit at all Chico Creek subwatershed streamflow gaging stations, this corresponding to a measurement objective function of 134.3

Parameter	Kitsap Creek	Wildcat Creek	Chico Creek (Taylor Road)	Dickerson Creek	Chico Creek (mainstream)
AGWETP	0.101	6.45E−04	3.60E−04	1.75E−03	1.38E−04
AGWRC	0.986	0.981	0.980	0.984	0.959
DEEPFR	2.80E−03	0.16	5.79E−03	9.87E−04	7.81E−03
INFILT	1.27	1.69E−01	2.90E−02	1.06E−01	3.71E−02
INTFW	1.00	3.01	5.73	3.65	5.02
IRC	0.72	0.65	0.85	0.76	0.85
LZETP	0.13	0.34	0.49	0.11	0.11
LZSN	5.0	12.5	32.7	13.8	24.5
UZSN	2.01	4.39	5.00	5.00	5.00

Traditional Tikhonov regularization was employed in the parameter estimation process.

Table 6 Weights calculated by PEST for regularization constraints applied during the adaptive regularization process

Parameter type	Regularization weight
AGWETP	1.408
AGWRC	1.093
DEEPFR	0.332
INFILT	1.337
INTFW	0.355
IRC	0.978
LZETP	0.628
LZSN	1.289
UZSN	1.009
IMP	1.762

Note that for all but the IMP parameters regularization constraints consisted of inter-subwatershed equality relationships. For the IMP parameters they were comprised of the assignment of preferred values. (By way of comparison, for unenhanced Tikhonov regularization a uniform weight of 2.37E−2 was applied for all parameter types.)

the local linear approximation necessary for computation of the Tikhonov regularization weight factor β^2. PEST-calculated weights for the 10 groups into which regularization constraints were subdivided (one for each estimated parameter type) are listed in Table 6. It is apparent from this table that regularization constraints were applied more strongly, and with greater discrimination between parameter types, during the adaptive regularization process than during the unenhanced Tikhonov regularization process.

Parameter estimation was next undertaken using truncated singular value decomposition (TSVD) — a subspace method — as a device for stabilizing the inverse problem. Using this methodology a solution to the inverse problem is sought within an orthogonal parameter subspace of reduced dimensionality to that of the original inverse problem, this reduction being such as to ensure that the condition number associated with the inverse problem rises no higher than a user-specified level; this value was set to 10^3 for the current case, thus ensuring numerical stability. See Doherty (2005) for further details of this methodology.

The TSVD inversion process resulted in a similar level of fit, and visually almost identical model-calculated flows, to those obtained using the other regularized inversion schemes discussed herein. However, estimated parameter values showed even greater inter-subwatershed variability and coincidence with upper and lower parameter bounds than those depicted in Tables 5a and 5b.

Calibration of an individual subwatershed model

A final calibration run was undertaken in which the Wildcat Creek subwatershed model was calibrated in isolation from other subwatershed models, with no regularization applied. It was found that stable numerical inversion could not be achieved unless the number of estimated parameters was reduced to 8. Hence only one IMP parameter was estimated (the others were tied to it such that they maintained a fixed ratio to it through the calibration process), and the relatively insensitive AGWETP and INTFW parameters were assigned values equal to those obtained through simultaneous hardwired subwatershed model calibration. Parameter values estimated through this process are depicted in Table 7. Model-to-measurement fit was visually almost identical to that displayed in Fig. 2. A Nash—Sutcliffe coefficient of 0.920 was obtained; this is similar to the value of 0.918 obtained for this same subwatershed when calibrated simultaneously with the other four subwatersheds using adaptive regularization. Objective function values were 25.8 for individual Wildcat Creek model calibration and 26.2 for adaptive regularization of this watershed model simultaneously with the other four subwatershed models. A slight deterioration in model-to-measurement fit was thus incurred through the adaptive regularization simultaneous calibration option. However, because the difference between these two objective functions is slight (less, in fact, than the objective function difference termination threshold for the simultaneous calibration option) no firm conclusions can be drawn from this exercise regarding relative performance of the two methods in minimizing model-to-measurement misfit. What is of interest, however, is that a comparison of Table 7 with the third column of Table 2a reveals that an almost identical level of fit between modeled and gaged mean daily flows at the Wildcat Creek gaging station can be achieved with significantly different sets of model parameters.

Table 7 Estimated parameter values for calibration of the Wildcat Creek subwatershed model alone

Parameter	Estimated value
IMP1	0.19
IMP2	0.23
IMP3	0.83
IMP4	0.11
AGWETP	1.15E−03
AGWRC	0.981
DEEPFR	0.18
INFILT	0.093
INTFW	1.88
IRC	0.67
LZETP	0.36
LZSN	14.4
UZSN	3.26

Discussion

The use of regularized inversion in model calibration brings with it many benefits. Principal among these is that it allows the parameter estimation process itself to introduce parsimony to model parameterization; normally this is done only to the extent necessary for the attainment of numerical stability, or to prevent ''over-fitting'' as defined by a suitable value for Φ_m. Because the inversion process is thereby freed from the imposition of ''pre-emptive parsimonizing'' as a necessary precursor to its deployment, it becomes maximally responsive to information contained within the calibration dataset in assigning values to estimated parameters. Model predictions are thus made with minimized error variance

(Moore and Doherty, 2005). For this reason, regularized inversion is employed as a matter of course in many scientific disciplines, including geophysics, medical image and signal processing and astrophysics (Haber, 1997). There is no reason why it should not find routine usage in the calibration of watershed models as well.

Formulation of the regularized inverse problem using the Tikhonov approach brings with it some advantages and some disadvantages. A distinct advantage is the fact that, theoretically at least, the modeler is able to exert a considerable influence on the direction taken by the parameter estimation process by specifying that his/her conception of parameter reasonableness be violated only to the minimum extent necessary to achieve an acceptable level of model-to-measurement fit. If the result of this process is then an unreasonable set of parameter values, this signifies that measurement noise is higher than anticipated, and a reduced level of model-to-measurement fit should thereby be requested in a repetition of the regularized parameter estimation process. Eventually it should be possible to obtain a set of parameter values that is hydrologically reasonable, while allowing the model to reproduce historical system behavior to the extent that this can be justified in the current modeling context.

Another benefit of the Tikhonov approach is that it allows the modeler to introduce ''soft data'' to the parameter estimation process. The Chico Creek example discussed above illustrates this process. Such data can lead not only to more realistic parameter values; it can also lead to a reduction in the degree of nonuniqueness associated with estimated parameter values, for it provides the means whereby certain, possibly large, subspaces of parameter space that lead to a low objective function, but are not necessarily in accord with a modeler's intuition, can be eliminated from further consideration.

Unfortunately, however, the advantages of the constrained minimization approach to regularized inversion cannot be realized unless the different regularization constraints introduced to the parameter estimation process are individually enforced to the greatest extent possible without compromising model-to-measurement fit. This is difficult to achieve when control over the strength with which these constraints are applied is in the hands of a single variable, this being β^2 of Eq. (5). Maximum efficacy of the Tikhonov method in achieving desirable parameter values requires that more controls be provided through which individual regularization constraints can be applied with different strengths, with these strengths depending on the freedom of movement granted to the parameters involved in these constraints by the calibration dataset. For those parameters, and/or parameter combinations, for which the calibration dataset is particularly uninformative, the constrained minimization calibration process must guarantee that user-supplied constraints are well respected. Only for those parameters that must vary in accordance with the demands of obtaining a good model-to-measurement fit, should regularization constraints be relaxed.

The adaptive regularization scheme described herein is an enhancement to the Tikhonov regularization methodology that, in many modeling contexts, including some that are of interest in watershed model calibration, allows this methodology to achieve its full potential of stable inversion with reasonable parameter outcomes. As a consequence, the watershed modeler has at his/her disposal a means of undertaking more sophisticated calibration than is available through traditional methodologies, allowing him/her to calculate parameter sets that better respect historical data on the one hand, while respecting notions of parameter reasonableness on the other. The latter is an outcome of the fact that the constrained minimization formulation of the inverse problem offers the modeler the ability to encapsulate his/her knowledge of the parameters that govern watershed processes as a set of constraints that will be respected by the inversion process insofar as this is possible, given the modeler's choice of desired goodness of fit. Because the adaptive regularization technique discussed herein enhances the ability of the constrained minimization process to respect those constraints, it thereby enhances the ability of the modeler to employ such knowledge as an integral part of the parameter estimation process.

Automated calibration of surface water models has been criticized for failing to take sufficient account of a modeler's intuitive knowledge of what constitutes a reasonable parameter set; see, for example, Lumb et al. (1994). In works such as this it is alleged that the ability of parameter estimation software to rapidly achieve a good fit between model outputs and field measurements seduces hydrologists into treating their models as ''black boxes'', thereby excluding their expert knowledge on parameter reasonableness from the calibration process. This is a criticism that the authors of this paper regularly face, in spite of the fact that computer assisted calibration of watershed models has been undertaken for over 30 years, has been the subject of intense research, and has won broad acceptance within many sectors of the watershed modeling community (Gupta et al., 2003). We share the opinion that the parameter estimation process suffers where modelers adopt an attitude of complacency toward parameter reasonableness in deference to an unstated assumption that ''a good fit is good enough''. It is hoped that the use of regularized inversion in general, and the adaptive enhancements to the Tikhonov scheme documented herein, will allow modelers to rapidly obtain both goodness of fit and parameter reasonableness where the modeling context allows this, or to better explore the tradeoffs between the two where it does not.

Software

PEST and supporting software are available free of charge from the following site: http://chl.erdc.usace.army.mil/pest.

Acknowledgement

The Puget Sound Naval Shipyard and Intermediate Maintenance Facility (PSNS & IMF) Environmental Investment (ENVVEST) Project supported this study. The authors would like to thank the PSNS & IMF ENVVEST Project participants for their contributions to this effort. Headquarters, US Army Corps of Engineers granted permission for publication of this paper.

References

Alley, W.M., Veenhuis, J.E., 1983. Effective impervious area in urban runoff modeling. Journal of Hydraulic Engineering 109 (2), 313–319.

Aster, R.C., Borchers, B., Thurber, C.H., 2005. Parameter Estimation and Inverse Problems. Elsevier, Amsterdam, 301 p.

Backus, G., Gilbert, F., 1967. Numerical application of a formalism for geophysical inverse problems. Geophysical Journal of the Royal Astronomical Society 13, 247–276.

Bates, B., Campbell, E., 2001. A Markov chain Monte Carlo scheme for parameter estimation and inference in conceptual rainfall-runoff modeling. Water Resources Research 37 (4), 937–947.

Bicknell, B.R., Imhoff, J.C., Kittle, J.L., Jobes, T.H., Donigian, A.S., 2001. HSPF User's Manual. Aqua Terra Consultants, Mountain View, California.

Box, G.E.P., Cox, D.R., 1964. An analysis of transformations. Journal of the Royal Statistical Society Series B 26, 211–243.

Box, G.E.P., Jenkins, G.M., 1976. Time Series Analysis: Forecasting and Control, rev. ed. Holden-Day, San Francisco, CA.

Boyle, D.P., Gupta, H.V., Sorooshian, S., 2000. Toward improved calibration of hydrologic models: combining the strengths of manual and automatic methods. Water Resources Research 36 (12), 3663–3674.

Bruckner, G., Handrock-Meyer, S., Langmach, H., 1998. An inverse problem from 2-D groundwater modeling. Inverse Problems 14 (4), 835–851.

Campbell, E.P., Bates, B.C., 2001. Regionalization of rainfall-runoff model parameters using Markov chain Monte Carlo samples. Water Resources Research 37 (3), 731–739.

Constable, S.C., Parker, R.L., Constable, C.G., 1987. Occam's inversion: a practical algorithm for generating smooth models from electromagnetic sounding data. Geophysics 52 (3), 289–300.

Conte, S.D., de Boor, C., 1972. Elementary Numerical Analysis. McGraw-Hill, New York.

deGroote-Hedlin, C., Constable, S., 1990. Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data. Geophysics 55 (12), 1624–1631.

Doherty, J., 2003a. Groundwater model calibration using pilot points and regularization. Ground Water 41 (2), 170–177.

Doherty, J., 2003b. PEST Surface Water Utilities. Watermark Numerical Computing, Brisbane, Australia.

Doherty, J., 2005. PEST: Model Independent Parameter Estimation: Fifth Edition of User Manual. Watermark Numerical Computing, Brisbane, Australia.

Doherty, J., Johnston, J.M., 2003. Methodologies for calibration and predictive analysis of a watershed model. Journal of the American Water Resources Association 39 (2), 251–265.

Duan, Q.S., Sorooshian, S., Gupta, V.K., 1992. Effective and efficient global optimization for conceptual rainfall runoff models. Water Resources Research 28 (4), 1015–1031.

Emsellem, Y., de Marsily, G., 1971. An automatic solution for the inverse problem. Water Resources Research 7 (5), 1264–1283.

Engl, H.W., Hanke, M., Neubauer, A., 1996. Regularization of Inverse Problems. Kluwer, Dordrecht.

ENVVEST Regulatory Working Group, 2002. FINAL Fecal Coliform Total Maximum Daily Load Study Plan for Sinclair and Dyes Inlets Quality Assurance Project Plan. October 4, 2002. Available from: <http://www.ecy.wa.gov/programs/wq/tmdl/watershed/sinclair-dyes_inlet/fc_tmdl_studyplan_final_draft_print.pdf>.

Gupta, H.V., Sorooshian, S., Hogue, T.S., Boyle, D.P., 2003. Advances in automatic calibration of watershed models. In: Duan, Q., Gupta, H., Sorooshian, S., Rousseau, A., Turcotte, R. (Eds.), Water Science and Application Series 6, 197–211.

Haber, E., 1997. Numerical strategies for the solution of inverse problems. PhD Thesis, University of British Columbia.

Haber, E., Ascher, U.M., Oldernburg, G., 2000. On optimization techniques for solving nonlinear inverse problems. Inverse Problems 16, 1263–1280.

Kuczera, G., 1983. Improved parameter inference in catchment models. 1. Evaluating parameter uncertainty. Water Resources Research 19 (5), 1151–1172.

Koren, V., 1993. The potential for improving lumped parameter models using remotely sensed data. In: Proceedings of the 13th Conference on Weather Analysis and Forecasting, August 2–6, 1993, Vienna, VA, USA, pp. 397–400.

Koren, V.I., Kuchment, L.S., 1971. A mathematical model of rain flood formation, optimization of its parameters and use in hydrological forecasting. Meteorology and Hydrology 11, 59–68 (in Russian).

Lumb, A.M., McCamon, R.B., Kittle Jr., J.L., 1994. Users manual for an expert system (HSPEXP) for calibration of the Hydrologic Simulation Program – Fortran: US Geological Survey Water-Resources Investigations Report 94-4168, 102 p.

Madsen, H., 2000. Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. Journal of Hydrology 235, 276–288.

Magette, W.L., Shanholtz, V.O., Carr, J.C., 1976. Estimating selected parameters for the Kentucky watershed model from watershed characteristics. Water Resources Research 12 (3), 472–476.

Menke, W., 1984. Geophysical Data Analysis: Discrete Inverse Theory. Academic Press, New York.

Moore, C., Doherty, J., 2005. The role of the calibration process in reducing model predictive error. Water Resources Research 41 (5), W05020.

Moore, C., Doherty, J., 2006. The cost of uniqueness in groundwater model calibration. Advances in Water Resources 29 (4), 605–623.

Portniaguine, O., Zhdanov, M.S., 1999. Focusing geophysical inversion images. Geophysics 64 (3), 874–887.

Schaake, J.C., Koren, V.I., Duan, Q.Y., 1996. A simple water balance model (SWB) for estimating runoff at different spatial and temporal scales. Journal of Geophysical Research, 101.

Skaggs, T.H., Kabala, Z.L., 1994. Recovering the release history of a groundwater contaminant. Water Resources Research 30 (1), 71–79.

Tikhonov, A., Arsenin, V., 1977. Solutions of Ill-posed Problems. V.H. Winston.

Tonkin, M.J., Doherty, J., 2005. A hybrid regularized inversion methodology for highly parameterized environmental models. Water Resources Research 41, W10412. doi:10.1029/2005WR003995.

USEPA, 2000. BASINS Technical Note 6: Estimating Hydrology and Hydraulic Parameters for HSPF. EPA-823-R-00-012.

van Loon, E.E., Troch, P.A., 2002. Tikhonov regularization as a tool for assimilating soil moisture data in distributed hydrological models. Hydrological Processes 16, 531–556.

Yokoo, Y., Kazama, M., Sawamoto, M., Nishimura, H., 2001. Regionalization of lumped water balance model parameters based on multiple regression. Journal of Hydrology 246, 209–222.